NRL Report 8322

Spectral Estimates of Band-Limited Signals

WILLIAM B. GORDON

Radar Analysis Staff
Radar Division

August 2, 1979

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NRL Report 8322	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): SPECTRAL ESTIMATES OF BAND-LIMITED SIGNALS	5. TYPE OF REPORT & PERIOD COVERED: Interim report on a continuing NRL problem
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): William B. Gordon	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Research Laboratory, Washington, D.C. 20375	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NRL Problem R12-51.101, Element 61153N, Project RR021-05-41
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Arlington, VA 22217	12. REPORT DATE: August 2, 1979
	13. NUMBER OF PAGES: 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Band-limited signals
Consistent spectral estimation
Nyquist sampling
Spectral analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We consider the problem of estimating the spectrum of a band-limited signal perturbed by additive white noise. Sharp bounds on the mean square errors of linear spectral estimates are computed and expressed as functions of time-bandwidth products and signal-to-noise power ratios. When the data are sampled at the Nyquist rate, *no* linear spectral estimator can produce significantly more accurate results than the conventional discrete Fourier transform. In the more general case, consistent spectral estimators can be obtained only if *both* the data rate *and* the length of the observation window are increased without bound.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

CONTENTS

UNCLASSIFIED

SPECTRAL ESTIMATES OF BAND-LIMITED SIGNALS

1. INTRODUCTION

A central problem in the theory of stationary time series is to estimate the spectrum of a stationary process $N(t)$ when the given data consist of samples of $s(t) = f(t) + N(t)$, where $f(t)$ is a deterministic trend of known functional form. In this paper we shall consider the dual problem: the functional form of $f(t)$ is unknown, the second-order statistics of $N(t)$ are known, and the problem is to estimate the spectrum of $f(t)$. The signals $f(t)$ will be assumed to be band-limited functions having continuous Fourier transforms $\hat{f}(\nu)$ which vanish for $|\nu| \geqslant W$. Such signals are pulse-like, and the amount of useful information contained in an observation time window depends as much on the position of the window as it does on its length. Accurate spectral estimation requires that the observation time window capture a significant portion of the signal energy.

We shall find that when the data are sampled at the Nyquist rate $2W$, consistent spectral estimators do not exist, in the sense that infinitely accurate estimates cannot be obtained from infinitely long data records. For signals with effective time duration T_e and signal-to-noise power ratio (S/N), most of the useful information is contained in a time window of length $N_c/(2W)$, where $N_c = (1/\pi)\sqrt{(2WT_e)^3(S/N)}$. For *any* linear spectral estimator there exist signals whose corresponding spectral estimates have relative mean square errors on the order of $1/N_c$ and absolute mean square errors which are almost as large as the largest produced by the conventional transform with N_c data points.

If the data are sampled at a rate higher than $2W$, longer time windows can be effectively used, and the conventional Fourier transform provides a consistent spectral estimator as the data rate increases without bound. If, however, the time window is fixed, consistent linear spectral estimators do not exist. Hence, to summarize, consistent linear spectral estimators exist only if *both* the data rate *and* the length of the time window increase without bound.

The problem of spectral estimation is essentially different from the problem of spectral peak detection and location, and hence our pessimistic results concerning the former do not preclude the possibility of high-resolution (supergain) spectral peak detectors, such as have been recently proposed for band-limited signals [1,2,3,4]. However, improved resolution necessarily results in a decrease in accuracy and detectability, and we hope to discuss this matter in subsequent papers.

Manuscript submitted April 16, 1979.

2. DEFINITIONS

2.1 The Class H(W)

Let $H(W)$ be the class of complex-valued band-limited functions $f = f(t)$ whose Fourier transforms $\hat{f} = \hat{f}(\nu)$ are continuous, vanish for $|\nu| \geqslant W$, and have L^2 derivatives defined almost everywhere. Then

$$f(t) = \int_{-W}^{+W} \hat{f}(\nu)\exp[2\pi i\nu t]\,d\nu, \tag{1}$$

$$\hat{f}(\nu) = \int_{-\infty}^{+\infty} f(t)\exp[-2\pi i\nu t]\,dt. \tag{2}$$

The functions f of class $H(W)$ are pulse-like in the sense that $|f(t)| \to 0$ as $|t| \to \infty$ (a consequence of the Riemann-Lebesque Lemma), and hence this class does not contain the sinusoidal functions whose spectra are delta functions (line spectra).

We follow the standard terminology of radar and communications theory [5,6] in defining the signal energy $E = E(f)$, mean time $\tilde{t} = \tilde{t}(f)$, and effective time duration $T_e = T_e(f)$ by

$$E = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-W}^{+W} |\hat{f}(\nu)|^2 d\nu, \tag{3}$$

$$\tilde{t} = (1/E) \int_{-\infty}^{+\infty} t|f(t)|^2 dt, \tag{4}$$

$$T_e^2 = (4\pi^2/E) \int_{-\infty}^{+\infty} (t-\tilde{t})^2 |f(t)|^2 dt. \tag{5}$$

We define a moment $M_2 = M_2(f)$ by

$$M_2 = 4\pi^2 \int_{-\infty}^{+\infty} t^2 |f(t)|^2 dt = \int_{-W}^{+W} |\hat{f}'(\nu)|^2 d\nu, \tag{6}$$

and we have the fundamental inequalities

$$|f(t)|^2 \leqslant \frac{W M_2(f)}{2\pi^2 t^2} \left\{ 1 - \frac{\sin^2(2\pi W t)}{(2\pi W t)^2} \right\}, \tag{7}$$

$$|\hat{f}(\nu)|^2 \leqslant \frac{1-(\nu/W)^2}{2} W M_2(f). \tag{8}$$

These two inequalities will be proved in Section 5.3, and they will be shown to be the sharpest possible on $H(W)$. The time-bandwidth product $2WT_e$ will appear frequently in our subsequent discussion, and we have the following "uncertainty relation":

$$2W T_e \geqslant \pi, \tag{9}$$

which can be shown to be the sharpest possible inequality of this type for $H(W)$.

Remark (1). In radar and communications theory, one has the uncertainty relation [6, p. 474]

$$B_e T_e \geqslant \pi,$$

where the effective bandwidth B_e is defined by

$$B_e^2 = (4\pi^2/E) \int (\nu - \widetilde{\nu})^2 |\hat{f}(\nu)|^2 d\nu,$$

with

$$\widetilde{\nu} = (1/E) \int \nu \, |\hat{f}(\nu)|^2 d\nu.$$

Using elementary inequalities one can easily show that on $H(W)$

$$|\widetilde{\nu}| \leqslant W, \; B_e \leqslant 2\pi W.$$

2.2 The Error Functions Q^2, $R^2(f)$, and R_o^2.

When the data are unperturbed by noise, $\hat{f}(\nu)$ will be estimated by discrete transforms of the type

$$\hat{f}_*(\nu) = \sum \overline{\beta}_n f(t_n),$$

and we shall derive inequalities of the type

$$|\hat{f}(\nu) - \hat{f}_*(\nu)|^2 \leqslant M_2(f) Q^2(\nu, \vec{t}, \vec{\beta}), \tag{11}$$

where $Q^2 = Q^2\,(\nu, \vec{t}, \vec{\beta})$ is a certain explicit function of the frequency ν, the sample point set $\vec{t} = \{t_1, t_2, \ldots, t_N\}$, and the weights $\vec{\beta} = \{\beta_1, \beta_2, \ldots, \beta_N\}$. These inequalities are *sharp* in the sense that for every given set of values for $\nu, \vec{t}, \vec{\beta}$, there exists a function f in $H(W)$ for which the inequality (11) becomes an equality. Hence the function Q^2 can be interpreted as the largest possible squared error in $\hat{f}_*(\nu)$ (when noise is absent) normalized by the moment M_2.

Suppose now that the data are perturbed by additive white noise $N(t)$ with zero mean. The data consist of samples of $s(t) = f(t) + N(t)$, and the problem is to estimate $\hat{f}(\nu)$ by sums of the type $\Sigma\, \overline{\beta}_n\, s(t_n)$. This estimate has mean and bias equal to $\hat{f}_*(\nu)$ and $[\hat{f}(\nu) - \hat{f}_*(\nu)]$, respectively, where $\hat{f}_*$ has the same meaning as before. The variance is independent of f and is given by

$$\text{Variance}\left\{\overline{\beta}_n s(t_n)\right\} = \sigma^2 \Sigma |\beta_n|^2, \tag{12}$$

where $\sigma^2 = \mathscr{E}\,[|N(t)|^2]$. For any f in $H(W)$, the mean square error in the spectral estimate $\Sigma\, \overline{\beta}_n\, s(t_n)$ is given by $R^2(f) = (\text{bias})^2 + \text{variance}$, or

$$R^2(f) = |\hat{f}(\nu) - \hat{f}_*(\nu)\,|^2 + \sigma^2 \Sigma\, |\beta_n|^2. \tag{13}$$

We define $R_o^2 = \sup\{R^2(f)\}$, where the supremum is taken over the set of all f in $H(W)$ having a given value of M_2. Then from the sharpness of the inequality (11), we have

$$R_o^2 = M_2\; Q^2(\nu, \vec{t}, \vec{\beta}) + \sigma^2 \Sigma |\beta_n|^2, \tag{14}$$

and for every $\epsilon > 0$ there exist f in $H(W)$ for which $R^2(f) \geqslant R_o^2 - \epsilon$.

2.3 The Conventional and Tapered Transforms

Explicit formulas for $Q^2(\nu, \vec{t}, \vec{\beta})$ for arbitrary values of ν, $\vec{t}$, and $\vec{\beta}$ will be derived in Section 5. For our present purposes we only need the result for the case when the data are equispaced and the weights $\vec{\beta}$ are conventional; i.e., $\beta_n = c_n$ where

$$c_n = \Delta t \exp[2\pi i \nu t_n] \tag{15}$$

and Δt is the data-point spacing. We write Q_c^2 for $Q^2\,(\nu, \vec{t}, \vec{c})$ and N_s for the number of sample points. When the sample point set $\{t_n\}$ is given by $t_n = n/(2W)$, where n varies over a set of integers $\{n: -N \leqslant n \leqslant N\}$, then $N_s = 2N{+}1$, and it turns out that Q_c^2 is independent of ν and is closely approximated by

$$Q_c^2 \simeq 2W/(\pi^2 N_s). \tag{16}$$

(The accuracy of this approximation increases with increasing N_s, and at $N_s = 3$ it is within 3% of the true value.)

For given values of ν and $\vec{t}$ there exists a unique set of weights $\vec{\beta} = \vec{\beta}(\nu, \vec{t})$ which minimizes $Q^2(\nu, \vec{t}, \vec{\beta})$, and hence might be said to be optimal in the absence of noise. These weights will hereafter be called the "tapered" weights because they vanish at $\nu = \pm W$ (a reflection of the fact that $\hat{f}(\pm W) = 0$ for all f in $H(W)$). The tapered weights are denoted by the symbol $\vec{b}$, and for ease of notation we write Q_b^2 for $Q^2(\nu, \vec{t}, \vec{b}(\nu, \vec{t}))$. Formulas for $\vec{b}$ are given in Section 5.5. Figure 1 shows Q_b^2/Q_c^2 as a function of ν/W for $N_s = 5, 11, 21$. This ratio is an even function of ν/W and the curves are only plotted for $\nu/W \geqslant 0$. Note that $Q_b^2/Q_c^2 = 0$ at $\nu = \pm W$, but the plunge to zero is very rapid, and $Q_b^2/Q_c^2 \geqslant 0.9$ over most of the frequency range. Also, the portion of the frequency range at which $Q_b^2/Q_c^2 \simeq 1$ increases as N_s increases.

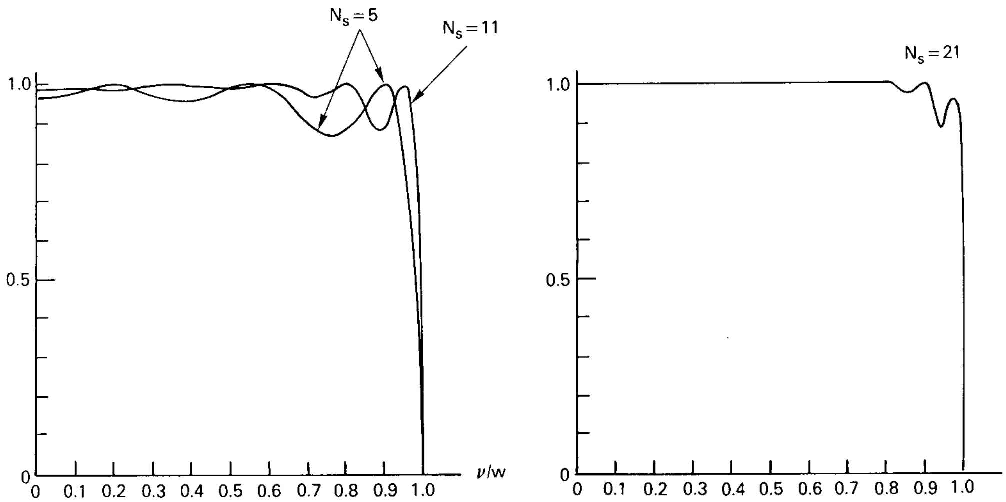

Fig. 1 — Q_b^2/Q_c^2 vs ν/W

Let V_c, V_b denote the variances of the spectral estimates $\Sigma\, \overline{c_n}\, s(t_n)$ and $\Sigma\, \overline{b_n}\, s(t_n)$, and let $\{t_n\}$ be as before. Then $|c_n| = 1/(2W)$, and from Eq. (12) we have

$$V_c = \sigma^2 N_s/(4W^2). \tag{17}$$

The ratio V_b/V_c turns out to be an even function of ν/W, and referring to Fig. 2 we note the following:

(i) The plunge to zero at $\nu = \pm W$ is again very rapid, and $V_b/V_c \geqslant 1$ over most of the frequency range.

(ii) The peak value of V_b/V_c decreases slowly with increasing N_s, and the peak locations converge to $\nu = \pm W$ as $N_s \to \infty$.

(iii) In Figs. 1 and 2 there is a close but not exact correspondence between the peaks of one curve and the dips of the other.

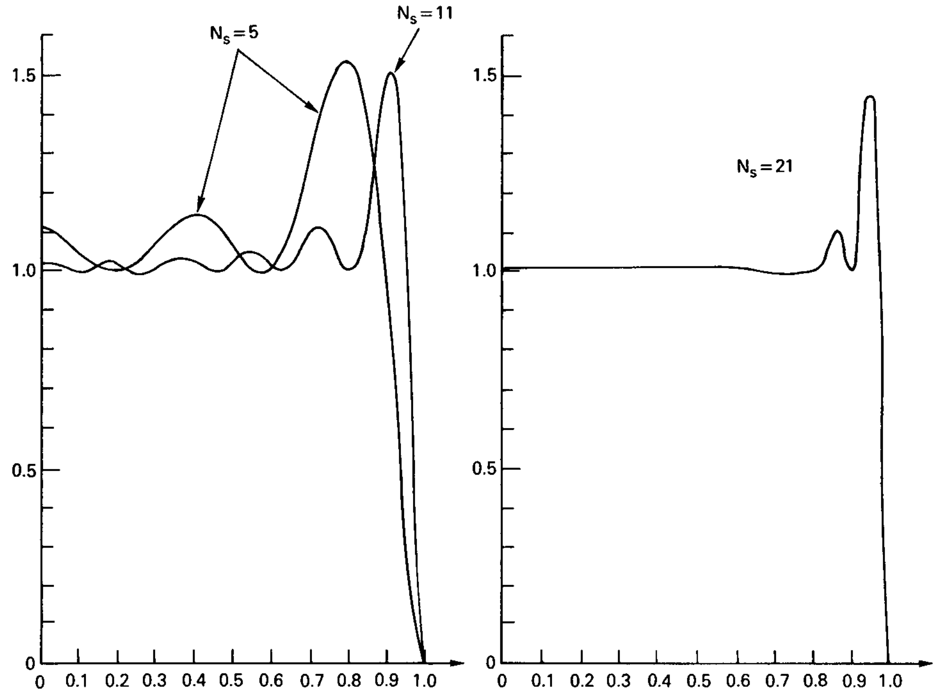

Fig. 2 — V_b/V_c vs ν/W

3. SAMPLING AT THE NYQUIST RATE

3.1 Preliminaries

Throughout this section it will always be assumed that the sample point set $\{t_n\}$ is given by $t_n = n/(2W)$, where the integer n varies over the set $\{n: -N \leqslant n \leqslant N\}$. As before, N_s will denote the number of samples ($= 2N + 1$), and T will denote the length of the time window. Hence $N_s \simeq 2WT$, and the time window is $[-T/2, T/2]$.

In particular, when we later consider the effects of expanding the time window, it is to be understood that the time window is to be prolonged in both directions. This is important, for this is a necessary condition for R_o^2 to converge to its infimum as $T \rightarrow \infty$, where the infimum is taken over the set of *all* linear spectral estimates $\Sigma\, \overline{\beta}_n s(t_n)$. The proof of this statement is tedious and will not be given; however, it is easy to see why it must be true because of the pulse-like nature of the functions of class $H(W)$.

UNCLASSIFIED

3.2 The Parameter N_c

We now assume that noise is present and that the conventional Fourier transform is used ($\beta_n = c_n$). Then from Eqs. (14) and (17) and the approximation (16) we have

$$R_o^2 = \frac{2WM_2}{\pi^2 N_s} + \frac{\sigma^2 N_s}{4W^2} . \tag{18}$$

Hence R_o^2 is a convex function of N_s, and its minimum value is attained at $N_s = N_c$, where

$$N_c = \left[\frac{8W^3 M_2}{\pi^2 \sigma^2}\right]^{1/2} . \tag{19}$$

For functions f satisfying $\tilde{t}(f) = 0$, Eq. (5) yields $M_2 = T_e^2 E$, and we have

$$N_c = (1/\pi) \left[(2WT_e)^3 (S/N) \right]^{1/2} , \tag{20}$$

where the signal-to-noise power ratio is defined by

$$(S/N) = \frac{\text{time - averaged signal power}}{\text{average noise power}} \text{, or}$$

$$(S/N) = \frac{E/T_e}{\sigma^2} . \tag{21}$$

The number N_c has another interesting property. When t is an integral multiple of $1/(2W)$, from Eq. (7) we have

$$|f(t)|^2 \leqslant \frac{WM_2}{2\pi^2 t^2} ,$$

and hence any f in $H(W)$ eventually becomes buried in noise; i.e., there exists a number T_o such that

$$|f(t)|^2 \leqslant \sigma^2 \text{ when } |t| \geqslant T_o .$$

Recalling that the data time window is $[-T/2, T/2]$, it is easy to verify that the condition $N_s \leqslant N_c$ is equivalent to the condition $T/2 \leqslant T_o$.

Although R_o^2 becomes an increasing function of N_s when $N_s > N_c$, we cannot assume that the same is true for $R^2(f)$ for any *particular* f. However, it can be shown that for any f in $H(W)$, $R^2(f)$ eventually becomes an increasing function of N_s, and using some gross estimates this can be shown to be the case when $N_s \geqslant N_c^2$. To prove this result, we consider the variation of the right-hand side of Eq. (13) when the value of N_s is increased by unity. The first term $|\hat{f}(\nu) - \hat{f}_*(\nu)|^2$ can change by no more than its largest possible value, viz.,

$M_2\, Q_c^2 = (2WM_2)/(\pi^2 N_s)$. Also, the second term changes (exactly) by the amount $\sigma^2/(4W^2)$. Hence the net change must be positive when $N_s \geqslant 8W^3 M_2/(\pi^2\sigma^2) = N_c^2$.

3.3 Nonexistence of Consistent Estimators

In the last paragraph we saw that when the conventional Fourier transform is used the error function R_o^2 eventually becomes an increasing function of N_s as $N_s \to \infty$. For each value of N_s we shall now choose a set of weights $\vec{\beta}$ which is "optimal" in the sense that it minimizes the right-hand side of Eq. (14) for given values of M_2 and σ^2. These weights will be called "σ-optimal," and when they are used, R_o^2 becomes a monotonically decreasing function of N_s. We define

$$R_\infty^2 = \lim_{N_s\to\infty} R_o^2 \tag{22}$$

and from the definitions it is evident that

$$R_\infty^2 = \inf\left\{R_o^2\right\} \tag{23}$$

where the infimum is taken over the set of *all* linear spectral estimates. Hence, for *any* linear spectral estimate $\Sigma\, \beta_n s(t_n)$ and for any $\delta > 0$ there exist functions f in $H(W)$ satisfying

$$R^2(f) > R_\infty^2 - \delta.$$

The calculation of R_∞^2 will be described in Section 5.5. It turns out that

$$R_\infty^2 = WM_2\, \frac{\cosh^2 [\pi N_c/2] - \cosh^2 [\pi N_c \nu/(2W)]}{(\pi N_c/2) \sinh \pi N_c}\ . \tag{24}$$

It follows that $R_\infty^2 > 0$, except when $\nu = \pm\, W$. Hence, consistent spectral estimators do not exist when the noise is white and the data are sampled at the Nyquist rate.

3.4 Conclusions

We shall now consider some consequences of Eq. (24). For notational ease we set

$$u = \pi N_c/2,\ \ \alpha = \nu/W. \tag{25}$$

Then

$$R_\infty^2 = WM_2\, \frac{\cosh^2 u - \cosh^2 \alpha u}{u \sinh 2u}\ . \tag{26}$$

Let X_c denote the value of R_o^2 when the conventional transform is used and $N_s = N_c$. From Eqs. (18) and (19) we get

$$X_c = \left[\frac{2M_2\sigma^2}{W\pi^2} \right]^{1/2}, \tag{27}$$

and therefore

$$R_\infty^2/X_c = (\pi/2) \frac{\cosh^2 u - \cosh^2 \alpha u}{\sinh 2u} . \tag{28}$$

When ν is close to $\pm W$ it is not unreasonable to compare R_∞^2 with the "trivial" estimate $\hat{f}(\nu) \equiv 0$. This estimate has no variance, and its mean square error X_o is given by the right-hand side of the (sharp) inequality (8). Hence,

$$R_\infty^2/X_o = \frac{2}{1-\alpha^2} \cdot \frac{\cosh^2 u - \cosh^2 \alpha u}{u \sinh 2u} . \tag{29}$$

The right-hand side of Eq. (28) attains its largest value when $\alpha = 0$, where it equals $(\pi/4)\tanh u$, and referring to Eqs. (9) and (20) one can easily verify that $\tanh u \simeq 1$ when $(S/N) \geqslant 1$.

The ratio R_∞^2/X_c decreases monotonically as u increases (i.e. the conventional Fourier transform improves as N_s increases). When $\nu = 0.9$ and $N_c = 1$, the ratio is 0.21. This means that for any linear spectral estimate there exist functions f in $H(W)$ for which $R^2(f)$ exceeds one-fifth the largest mean-square error produced by a conventional one-point transform (which occurs at some function different from f).

We now examine the right-hand side of Eq. (29). This ratio converges to unity as $\nu \to \pm W$, and its minimum value is attained at $\nu = 0$, where it equals $(1/u) \tanh u \simeq 1/u = 2/(\pi N_c)$. We therefore draw the following conclusions:

Conclusion 1.

When the data are sampled at the Nyquist rate, every linear spectral estimator produces mean square errors having the same order of magnitude as the *largest* produced by a conventional Fourier transform with N_c data points, except near $\nu = \pm W$, where the errors have the same order of magnitude as those produced by the trivial estimate $\hat{f}(\nu) \equiv 0$.

The ratio R_∞^2/X_o can also be interpreted as a bound on the relative errors $R^2(f)/|\hat{f}(\nu)|^2$. For there exist f for which $R^2(f) > R_\infty^2$, and for any f we have $|\hat{f}(\nu)|^2 \leqslant X_o$ since X_o is the right-hand side of Eq. (8). Hence, there are always f for which $R^2(f)/|\hat{f}(\nu)|^2$ exceeds R_∞^2/X_o, and we are led to the following:

Conclusion 2.

When the data are sampled at the Nyquist rate, every linear spectral estimator produces relative mean square errors which are on the order of $1/N_c$ near $\nu = 0$ and unity near $\nu = \pm W$.

4. OVERSAMPLING

4.1 The Case of Variable T

In this section we consider the effects of oversampling. That is, we now suppose that functions of class $H(W)$ are sampled at a rate $2kW$, $k \geqslant 1$. As before, we assume that the time window is of the form $[-T/2, T/2]$, so that the window is prolonged in *both* directions when T is increased. We also assume that the noise remains white, which is usually the case for receiver and thermal noise.

The results of Section 3 are based on certain closed-form expressions for Q_b^2 and R_∞^2 which are derived in Section 5 for the case of sampling at the Nyquist rate. These derivations require the closed-form inversion of certain matrices, which, unfortunately, we have been unable to effect for the case of oversampling. Hence we are presently unable to give a quantitative description of how much useful information is contained in a fixed time window when the data rate is increased without bound. However, it can be shown that consistent spectral estimators exist only if *both* the data rate $2kW$ *and* the length T of the time window are allowed to increase without bound. Results of this type are apparently known, and our discussion will be brief. (Cf. the discussion in Section B.1 of Blackman and Tukey [7] which suggests the existence of results of this nature.)

The class $H(W)$ can be considered as a subset of $H(kW)$, and error bounds can be obtained by replacing each occurrence of W with kW in the formulas of Section 3. The error bounds thus obtained are sharp for the class $H(kW)$, but are not sharp for the class $H(W)$, and we shall refer to this method as the "inexact" method.

Recall that the error function R_o^2 of the conventional transform attains its minimum value X_c when $N_s = N_c$, or equivalently, when $T = N_c/(2W)$. One effect of increasing the data rate is to increase the value of N_c and hence permit the use of longer data records. For by applying the inexact method to Eq. (19), we see that N_c increases by a factor of $k^{3/2}$ when W is replaced with kW. Moreover, from Eq. (27), X_c decreases by the factor $1/\sqrt{k}$. Hence, conventional transform provides a consistent spectral estimator if the length of the time window is increased by a factor of $\sqrt{k}$ when the sampled rate is increased to $2kW$ and $k \rightarrow \infty$.

4.2 The Case of Fixed T

We shall now suppose that T is fixed, and show that consistent spectral estimators do not exist when $k \rightarrow \infty$. The proof uses Shannon's Sampling Theorem, according to which every band-limited function f whose Fourier transform is of class L^2 has the representation

$$f(t) = \sum_{n=-\infty}^{+\infty} f_n \frac{\sin[\pi(n-2Wt)]}{[\pi(n-2Wt)]}, \tag{30}$$

where we set $f_n = f[n/(2W)]$. The bi-infinite sequence $\{f_n\}$ belongs to the class l_2; i.e.,

$$\sum_{-\infty}^{+\infty} |f_n|^2 < \infty.$$

For each $k > 1$ the data $s(t) = n(t) + N(t)$ are sampled in the window $[-T/2, T/2]$ at the points $t = n/(2kW)$, where the integers n are bounded by

$$|n| \leqslant N_k = \text{integral part of } (kWT).$$

Let $\vec{\beta}^{(k)} = \{\beta_n^{(k)}\}$, $(-N_k \leqslant n \leqslant N_k)$ be a weight vector which produces the spectral estimate $\Sigma\, \overline{\beta}_n^{(k)} s\,[n/(2kW)]$. We shall suppose that the $\vec{\beta}^{(k)}$ can be chosen to make the mean square error converge to zero as $k \to \infty$, and derive a contradiction. Setting $t = n/(2kW)$ in Eq. (30) we get the formal result

$$\sum_{m=-N_k}^{N_k} \overline{\beta}_m^{(k)} s\left[\frac{m}{2kW}\right] = \sum_{m=-N_k}^{N_k} \overline{\beta}_m^{(k)} f\left[\frac{m}{2kW}\right] + \sum_{m=-N_k}^{N_k} \overline{\beta}_m^{(k)} N\left[\frac{m}{2kW}\right]. \tag{31}$$

Using Eq. (30), the first sum of the right-hand side can be expressed as an infinite series involving the functional values $f_n = f[n/(2W)]$, and this series can easily be shown to converge if the weight vectors $\vec{\beta}^{(k)}$ are chosen to make the mean error converge to zero as $k \to \infty$. The second sum can be expressed as a finite linear combination of the values $N^{(k)}[t/(2W)]$, where the noise process $N^{(k)}(t) = N(t/k)$ is assumed to be white. Hence every estimator of type (31) is a linear combination of functional values of f and white noise spaced at $\Delta t = 1/(2W)$, and therefore, from the results of the previous section, cannot be consistent.

5. MATHEMATICAL DERIVATIONS

5.1 General Theory

We first consider the noise-free case, and describe the mathematical tools used in the calculation of the error function Q^2. Referring to Eq. (10) for notation, it is now convenient to define the function $Q^2 = Q^2(\nu, \vec{t}, \vec{\beta})$ by

$$Q^2 = \sup_f \left\{ \frac{|\hat{f}(\nu) - \hat{f}_*(\nu)|^2}{||f||^2} \right\}, \tag{32}$$

where $||\cdot||$ is a norm defined on a Hilbert function space H, and the supremum is taken over all f belonging to H. Hence

$$|\hat{f}(\nu) - \hat{f}_*(\nu)|^2 \leqslant ||f||^2 \, Q^2(\nu, \vec{t}, \vec{\beta}), \tag{33}$$

and from Hilbert space generalities this inequality is sharp. We shall subsequently take $H = H(W)$ and $||f||^2 = M_2(f)$, so that the sharpness of Eq. (33) implies that Eq. (11) is sharp, but we shall first discuss the theory in more general terms.

In order for the right-hand side of Eq. (32) to be finite, it is necessary that the norm $||\cdot||$ satisfy the following two conditions:

Condition 1. For each t, the evaluation map $f \to f(t)$ is continuous with respect to $||\cdot||$, and

Condition 2. For each ν, the map $f \to \hat{f}(\nu)$ is continuous.

Hilbert function spaces H satisfying Condition 1 are called reproducing kernel Hilbert spaces [8], and for such spaces the Riesz Representation Theorem guarantees the existence of functions $K_t = K_t(s)$ satisfying

$$f(t) = (f, K_t), \text{ for all } f \text{ in } H, \tag{34}$$

where $(\cdot,\cdot)$ denotes the inner product corresponding to the norm $||\cdot||$. The Hermitian matrix

$$K_t(s) = (K_t, K_s) \tag{35}$$

is called the reproducing kernel. Similarly, when Condition 2 is satisfied there exist functions $e_\nu = e_\nu(t)$ in H defined by

$$\hat{f}(\nu) = (f, e_\nu). \tag{36}$$

We can now write Eq. (32) in the form

$$Q^2 = \sup_f \left\{ \frac{|(f, e_\nu - \Sigma\beta_n K_{t_n})|^2}{||f||^2} \right\},$$

and from Schwarz's inequality it follows that

$$Q^2 = ||e_\nu - \Sigma\beta_n K_{t_n}||^2. \tag{37}$$

Letting $<\cdot,\cdot>$ denote the standard inner product on C^N and expanding Eq. (37), we get

$$Q^2 = < K\vec{\beta}, \vec{\beta} > - < \vec{\beta}, \vec{J} > - < \vec{J}, \vec{\beta} > + ||e_\nu||^2, \tag{38}$$

where the matrix $K = K(\vec{t})$ and the vector $\vec{J} = \vec{J}(\nu, \vec{t})$ are defined by

$$K_{nm} = (K_{t_m}, K_{t_n}) = K_{t_m}(t_n), \tag{39a}$$

$$J_n = (e_\nu, K_{t_n}) = e_\nu(t_n). \tag{39b}$$

Equations (38) and (39) are the basic equations of this theory, and from them we see that the error function Q^2 is a nonhomogeneous polynomial in $\vec{\beta}$. With a little calculus it is easy to show that the "optimum" value of $\vec{\beta}$ which minimizes Q^2 is given by $\vec{\beta} = \vec{b}$, where $\vec{b} = \vec{b}(\nu, \vec{t})$ is given by

$$\vec{b} = K^{-1}\vec{J}. \tag{40}$$

The corresponding minimum value of Q^2 is denoted by Q_b^2, and substituting Eq. (40) into Eq. (38) we get

$$Q_b^2 = \|e_\nu\|^2 - \langle K^{-1}\vec{J}, \vec{J}\rangle = \|e_\nu\|^2 - \langle \vec{b}, \vec{J}\rangle. \tag{41}$$

When H = $H(W)$ the weight vectors $\vec{b}$ are the tapered weights discussed in Section 2. We shall later see that σ-optimal weights used in calculation of R_∞^2 are obtained by inverting the matrix $K + (\sigma^2/M_2)I$.

5.2 The Space H(W)

It is well known that the ordinary L^2 norm satisfies Condition 1 on the class of band-limited functions f having the representation (1), and for a neat account of the L^2 theory of linear operators, we refer the reader to Ref. 9. However, it is a standard exercise to show that the L^2 norm does not satisfy Condition 2, and we therefore define an inner product $(\cdot,\cdot)$ and corresponding norm $\|\cdot\|$ by

$$(f,g) = \int_{-W}^{+W} D\hat{f}(\nu) \cdot \overline{D\hat{g}(\nu)}\, d\nu, \quad \|f\|^2 = (f,f), \tag{42}$$

where D denotes the differentiation operator. From Eq. (6) we have

$$\|f\|^2 = M_2(f). \tag{43}$$

Let $C_o^\infty(W)$ denote the class of band-limited functions $f = f(t)$ whose Fourier transforms $\hat{f} = \hat{f}(\nu)$ are of class C^∞ and vanish for $|\nu| \geqslant W$. It is now convenient to define $H(W)$ as the completion of $C_o^\infty(W)$ with respect to the norm $\|\cdot\|$. For smooth functions $\hat{f} = \hat{f}(\nu)$ we have

$$\hat{f}(\nu_1) - \hat{f}(\nu_0) = \int_{\nu_0}^{\nu_1} D\hat{f}(\nu)\, d\nu;$$

hence, from Schwarz's inequality we have

$$|\hat{f}(\nu_1) - \hat{f}(\nu_0)|^2 \leqslant |\nu_1 - \nu_0| \int_{\nu_0}^{\nu_1} |D\hat{f}(\nu)|^2 \, d\nu,$$

and therefore

$$|\hat{f}(\nu_1) - \hat{f}(\nu_0)| \leqslant \sqrt{|\nu_1 - \nu_0|} \;\; \| f \|.$$

From this inequality it is easily shown that the norm $\|\cdot\|$ satisfies Conditions 1 and 2, and that the transforms $\hat{f}$ of functions f in $H(W)$ are absolutely continuous and vanish at the end points $\nu = \pm W$. The transforms $\hat{f}$ need not be differentiable at every point, and if the transforms $\hat{f}, \hat{g}$ which appear in Eq. (42) are piecewise smooth, then one can evaluate the integral by integrating on each piece and taking the sum. More generally, one can replace $\hat{f}, \hat{g}$ by the appropriate Cauchy sequences in $C_o^\infty(W)$ and take the limit as $n \to \infty$.

Remark 2. Function spaces whose norms incorporate L^p norms of kth order derivatives are called Sobolev spaces, and are generally denoted by the symbols L_k^p, $W_{p,k}$, and H_k for the case $p = 2$. They are currently much in use in differential equations and in the calculus of variations, and we refer the reader to Refs. 10, 11, and 12 for the linear theory and to Ref. 13 for the nonlinear theory. Condition 2 for the norm of $H(W)$ is a particular example of Sobolev Embedding Theorem. Sobolev space techniques have also been used by Sobolev and his students in studies of efficient numerical quadrature (cubature) on high-dimensional manifolds, and we refer the reader to the expository paper by Haber [14] for an account of this subject.

5.3 Basic Formulas For H(W)

When $H = H(W)$, the functions $Q^2 = Q^2(\nu, \vec{t}, \vec{\beta})$ can be computed for general $\nu, \vec{t}, \vec{\beta}$ by means of Eqs. (38), (39), and the following identities:

$$K_s(t) = \frac{1}{4\pi^3} \left\{ \frac{\sin[2\pi W(t-s)]}{st(t-s)} - \frac{\sin(2\pi Ws)\sin(2\pi Wt)}{2\pi W\, s^2 t^2} \right\}, \; (s, t, s-t \neq 0), \qquad (44a)$$

$$K_t(t) = \frac{1}{4\pi^3 t^3} \left\{ 2\pi Wt - \frac{\sin^2(2\pi Wt)}{2\pi Wt} \right\}, \qquad (t \neq 0), \qquad (44b)$$

$$K_0(t) = K_t(0) = \frac{1}{4\pi^3 t^3} \left\{ \sin(2\pi Wt) - (2\pi Wt)\cos(2\pi Wt) \right\}, \qquad (t \neq 0), \qquad (44c)$$

$$K_0(0) = \frac{2}{3} W^3, \qquad (44d)$$

$$e_\nu(t) = \frac{1}{4\pi^2 t^2}\left\{ e^{2\pi i\nu t} - \cos(2\pi Wt) - (i\nu/W)\sin(2\pi Wt)\right\}, \quad (t \neq 0), \quad (45a)$$

$$e_\nu(0) = (1/2)\,(W^2 - \nu^2), \quad (45b)$$

$$\| e_\nu \|^2 = (W^2 - \nu^2)/(2W). \quad (45c)$$

When the data are sampled at the Nyquist rate $2W$, $t_n = n/(2W)$, and Eqs. (39) specialize to

$$\begin{cases} K_{mn} = 0, & (m \neq n, m \neq 0, n \neq 0), \\ K_{mm} = (2W^3)/(\pi^2 m^2), & (m \neq 0), \\ K_{mo} = (-1)^{m+1}(2W^3)/(\pi^2 m^2), & (m \neq 0), \\ K_{oo} = (2/3)W^3, & \end{cases} \quad (46)$$

$$\begin{cases} J_m = [W^2/(\pi^2 m^2)]\,[\exp(\pi i m\nu/W) - (-1)^m], & (m \neq 0), \\ J_o = (1/2)\,(W^2 - \nu^2), & \end{cases} \quad (47)$$

$$\| e_\nu \|^2 = (W^2 - \nu^2)/(2W). \quad (48)$$

Remark 3. The basic inequalities (7) and (8) are direct consequences of Eqs. (44b) and (45c). For, using the Schwarz inequality we get

$$|f(t)|^2 = |(f,K_t)|^2 \leqslant \|f\|^2\,\|K_t\|^2 = \|f\|^2\,K_t(t),$$

$$|\hat{f}(\nu)|^2 = |(f,e_\nu)|^2 \leqslant \|f\|^2\,\|e_\nu\|^2.$$

Proofs. Integrating Eq. (34) by parts we obtain

$$f(t) = (f,K_t) = -\int_{-W}^{+W} \hat{f}(\nu) \cdot \overline{D^2 \hat{K}_t(\nu)}\,d\nu$$

Hence, comparing this equation with Eq. (1) we see that $\hat{K}_t = \hat{K}_t(\nu)$ is the solution to

$$\left\{ -D^2 \hat{K}_t(\nu) = \exp[-2\pi i\nu t];\ \hat{K}_t(-W) = \hat{K}_t(W) = 0 \right\}.$$

(The boundary conditions $\hat{K}_t(\pm W) = 0$ are required because K_t belongs to the class $H(W)$, whose transforms vanish at the endpoints $\nu = \pm W$.) Hence, for $t \neq 0$ we have

$$\begin{cases} \hat{K}_t(\nu) = \dfrac{1}{4\pi^2 t^2} \left\{ e^{-2\pi i \nu t} - \cos(2\pi W t) + (i\nu/W)\sin(2\pi W t) \right\}, \\ \hat{K}_0(\nu) = (1/2)(W^2 - \nu^2). \end{cases} \tag{49}$$

To obtain Eqs. (44), compute $K_s(t) = (\hat{K}_s, \hat{K}_t)$ directly from Eqs. (42) and (49).

To obtain Eqs. (45), we have

$$e_\nu(t) = (e_\nu, K_t) = \overline{(K_t, e_\nu)} = \overline{\hat{K}_t(\nu)},$$

and we use Eq. (49). It is interesting to note that $\hat{e}_{\nu_o}(\nu)$ is piecewise smooth, with

$$e_{\nu_o}(\nu) = [1/(2W)] \left\{ (W^2 - \nu_o^2) - (\nu - \nu_o)(W\, h(\nu - \nu_o) + \nu_o) \right\}, \tag{50}$$

where h is the Heaviside function, $h(\nu) = \text{sign}(\nu)$.

5.4 Derivation of Formulas for Q_c^2

Let the data-point set $\{t_n\}$ be given by $t_n = n/(2W)$ where n varies over the set of $N_s = (2N + 1)$ integers $\{n: -N \leqslant n \leqslant N\}$. The conventional weights are given by $\beta_n = c_n$ where

$$c_n = (1/2W) \exp(\pi i n \nu / W), \tag{51}$$

and by direct substitution into Eqs. (38), (46), (47), and (48) we get

$$Q_c^2 = (W/2)\left\{(1/3) - (2/\pi^2) \sum_1^N (1/n^2)\right\}. \tag{52}$$

We shall have occasion to use the identities

$$\sum_1^\infty \frac{\cos 2\pi n t}{n^2} = \pi^2 (t^2 - t - \frac{1}{6}), \quad 0 \leqslant t \leqslant 1, \tag{53a}$$

$$\sum_{-\infty}^{+\infty} \frac{\exp 2\pi i n t}{1 + \gamma^2 n^2} = \frac{\pi}{\gamma} \frac{\exp 2\pi t/\gamma + \exp 2\pi(1-t)/\gamma}{\exp 2\pi/\gamma - 1}, \tag{53b}$$

whose proofs consist in verifying that the left-hand sides are the Fourier series of the right-hand sides. Setting $t = 0$ in Eq. (53a) we get the well known identity $\sum_1^\infty (1/n^2) = 1/6$. Hence,

$$Q_c^2 = (W/\pi^2) \sum_{N+1}^{\infty} (1/n^2),$$

and Eq. (16) is obtained using the approximation

$$\sum_{N+1}^{\infty} (1/n^2) \simeq \int_{N+1}^{\infty} \frac{dx}{x^2} = \frac{1}{N+1} \simeq \frac{2}{N_s}.$$

5.5 Derivation of Formulas for $\mathbf{R}_\infty^2$

We shall now sketch the calculation of the σ-optimal weights defined in Section 3.3. These are the weights that minimize R_o^2 for given values of M_2 and σ^2, and referring to Eq. (14) we see that the tapered weights (defined in Section 2.3) are obtained by setting $\sigma = 0$. We shall assume $t_n = n/(2W)$ where n varies over a set L of integers, and that the integer 0 belongs to this set. (However, the set is not assumed to be symmetric about 0.)

Referring to Eqs. (14) and (38), we have

$$R_0^2/M_2 = \langle (K + \frac{\sigma^2}{M_2} I)\, \vec{\beta}, \vec{\beta} \rangle - \langle \vec{\beta}, \vec{J} \rangle - \langle \vec{J}, \vec{\beta} \rangle + \|e_\nu\|^2.$$

Hence the minimum occurs when $(K + (\sigma^2/M_2)\, I)\, \vec{\beta} = \vec{J}$, and from Eq. (46) this equation reduces to the system

$$\begin{cases} K_{mo}\,\beta_o + (K_{mm} + \dfrac{\sigma^2}{M_2})\,\beta_m = J_m, & (m \neq 0), \\[2ex] (K_{oo} + \dfrac{\sigma^2}{M_2})\,\beta_o + \Sigma_L' K_{on}\beta_n = J_o, \end{cases}$$

where Σ_L' denotes summation over the nonzero integers n in L. The solution is readily obtained by solving first for β_m and then for β_o. Let the quantities α, a, A, B, C be defined by

$$\alpha = \nu/W, \tag{54}$$

$$a = \left[\frac{\sigma^2}{2W^3 M_2}\right]^{1/2} = 2/(\pi N_c),$$

$$\left\{\begin{aligned} A &= \frac{1-\alpha^2}{2} - \frac{1}{\pi^2}\sum\nolimits_L' \frac{1-(-1)^n\cos\pi n\alpha}{n^2(1+a^2\pi^2n^2)}, \\ B &= (\frac{1}{3}+a^2) - \frac{1}{\pi^2}\sum\nolimits_L' \frac{1}{n^2(1+a^2\pi^2n^2)}, \\ C &= \frac{1}{\pi^2}\sum\nolimits_L' \frac{(-1)^n\sin\pi n\alpha}{n^2(1+a^2\pi^2n^2)}. \end{aligned}\right. \tag{55}$$

Then the σ-optimal weights $\vec{\beta}$ are given by

$$\left\{\begin{aligned} \beta_o &= (A - iC)/(2WB), \\ \beta_n &= c_n + (-1)^n(\beta_o - \frac{1}{2W}), \end{aligned}\right. \tag{56}$$

and the corresponding minimum value of R_o^2 is given by

$$R_0^2 = WM_2\left\{A - \frac{A^2+C^2}{2B}\right\}. \tag{57}$$

The value of R_∞^2 is computed according to Eq. (22), and therefore each of the sums in Eqs. (55) becomes an infinite series: $\Sigma_L' = \Sigma_{n<o} + \Sigma_{n>o}$. Then $C = 0$, and each of the infinite series in the expressions for A and B is decomposed into the sum of two infinite series by means of the identity

$$\frac{1}{n^2(1+a^2\pi^2n^2)} = \frac{1}{n^2} - \frac{a^2\pi^2}{1+a^2\pi^2n^2}.$$

Closed-form expressions for the resulting series are then obtained by means of Eq. (53b).

As already mentioned, expressions for the tapered weights are obtained from Eq. (56) by setting $a = 0$ in Eq. (55), and the right-hand side of Eq. (57) then represents $W\,Q_b^2$.

Finally, in the derivations we have assumed that the data-point set contains zero. When this is not the case, K is a diagonal matrix and the derivations are even simpler. In particular, we get

$$R_o^2 = WM_2\,A.$$

REFERENCES

1. J.A. Cadzow, "Improved Spectral Estimation From Incomplete Sampled Data Observations," RADC Spectrum Analysis Workshop, Rome Air Development Center, Rome, N.Y. 85-96 (May 1978).
2. J.A. Cadzow, "An Extrapolation Procedure for Band-Limited Signals," IEEE Trans. Acoust. Speech Signal Process. ASSP-27, 4-12 (1979).
3. A. Papoulis, "A New Algorithm in Spectral Analysis and Band-Limited Extrapolation," IEEE Trans. Circuits Syst. CAS-22, 735-742 (1975).
4. A. Papoulis and C. Chamzas, "Adaptive Extrapolation and Hidden Periodicities," RADC Spectrum Analysis Workshop, Rome Air Development Center, Rome, N.Y., 85-96 (May 1978).
5. G.W. Deley, "Waveform Design," in *Radar Handbook* (M.I. Skolnik, ed.), McGraw-Hill, New York, N.Y. 1970.
6. M.I. Skolnik, *Introduction to Radar Systems*, McGraw-Hill, New York, N.Y., 1962.
7. R.B. Blackman and J.W. Tukey, *The Measurement of Power Spectra*, Dover, New York, N.Y. 1959.
8. N. Aronszajn, "Theory of Reproducing Kernels," Trans. Am. Math. Soc. 68, 337-404 (1950).
9. A. Papoulis, "Limits on Bandlimited Signals," Proc. IEEE 55, 1677-1686 (1967).
10. J. Barros-Neto, *An Introduction to the Theory of Distributions*, Marcel Dekker, New York, N.Y. 1973.
11. L. Bers, F. John, and M. Schechter, *Partial Differential Equations*, Interscience, New York, N.Y. 1964.
12. F. Treves, *Basic Linear Partial Differential Equations*, Academic Press, New York, N.Y. 1975.
13. R.S. Palais, *Foundations of Global Non-Linear Analysis*, Benjamin, New York, N.Y. 1968.
14. S. Haber, "Numerical Evaluation of Multiple Integrals," SIAM Rev. 12, 481-526 (1970).